INFECTION AND IMMUNITY, Oct. 2004, p. 5613–5621 Vol. 72, No. 10
0019-9567/04/$08.00+0 DOI: 10.1128/IAI.72.10.5613–5621.2004

Structural Organization of the pFra Virulence-Associated Plasmid of Rhamnose-Positive *Yersinia pestis*

Andrey Golubov,[1] Heinrich Neubauer,[2] Christina Nölting,[1] Jürgen Heesemann,[1] and Alexander Rakin[1]*

Max von Pettenkofer Institute of Hygiene and Medical Microbiology[1] and Institute for Microbiology, German Federal Armed Forces,[2] Munich, Germany

Received 19 May 2004/Returned for modification 23 June 2004/Accepted 12 July 2004

The 137,036-bp plasmid pG8786 from rhamnose-positive *Yersinia pestis* G8786 isolated from the high mountainous Caucasian plague focus in Georgia is an enlarged form of the pFra virulence-associated plasmid containing genes for synthesis of the antigen fraction 1 and phospholipase D. In addition to the completely conserved genes of the pFra backbone, pG8786 contains two large regions consisting of 4,642 and 32,617 bp, designated regions 1 and 2, respectively. Region 1 retains a larger part of *Salmonella enterica* serovar Typhi plasmid pHCM2 resembling the backbone of pFra replicons, while region 2 contains 25 open reading frames with high levels of similarity to the transfer genes of the F-like plasmids. Surprisingly, region 1 is also present in the pFra plasmid of avirulent *Y. pestis* strain 91001 isolated in Inner Mongolia, People's Republic of China. Despite the fact that some genes typically involved in conjugative transfer of the F-like replicons are missing in pG8786, we cannot exclude the possibility that pG8786 might be transmissive under certain conditions. pG8786 seems to be an ancient form of the pFra group of plasmids that were conserved due to the strict geographical isolation of rhamnose-positive *Y. pestis* strains in the high mountainous Caucasian plague locus.

Yersinia pestis, the causative agent of plague, is thought to be a recently emerged pathotype of the enteropathogen *Yersinia pseudotuberculosis* (1, 26). Plague is a zoonosis. Transmission of *Y. pestis* by a fleabite usually causes bubonic plague. Further dissemination of bacteria through the bloodstream leads to secondary septicemic plague. Dissemination into the lungs can cause a more contagious secondary pneumonia. Three *Y. pestis* biovars, *Y. pestis* bv. Antiqua, *Y. pestis* bv. Mediaevalis, and *Y. pestis* bv. Orientalis, which are believed to be the causative agents of the historical plague pandemics, are distinguished by the ability to ferment glycerol and the nitrification activity (7). However, in addition to these organisms there is a group of *Y. pestis* isolates distributed in various countries of the former USSR, Mongolia, People's Republic of China, and Morocco that share certain characteristics with the closely related species *Y. pseudotuberculosis* (2). These isolates ferment rhamnose, are also dependent on additional nutrients, and exhibit elective virulence (they are less virulent in guinea pigs but highly virulent in mice). These strains are described in the literature as causes of occasional human or animal plague cases, but they have rarely been associated with epizootics of plague (2, 28). To separate these rhamnose-positive isolates from the main group of *Y. pestis* strains, it has been proposed that they should be named *Yersinia pestoides* or Pestoides (18). Alternatively, they were named on the basis of the places where they were first isolated (i.e., *Y. pestis* subsp. *caucasica*, *Y. pestis* subsp. *ulegeica*, *Y. pestis* subsp. *altaica*, etc.) (2).

The main acquisitions of the plague microbe thought to be responsible for its virulence are two plasmids. pPla (also designated pYP, pPCP1, or pPst) encodes the plasminogen activator and the bacteriocin pesticin. pFra (also designated pMT1 or pYT) is responsible for the synthesis of fraction 1 antigen and phospholipase D. The plasminogen activator is involved in the dissemination of the plague bacterium from the site of the initial fleabite, while phospholipase D (previously accepted as a murine toxin) plays a major role in survival of plague bacteria in fleas (12). All pathogenic yersiniae contain the virulence-associated pYV plasmid, which encodes finely tuned type III secretion machinery consisting of anti-phagocytic factors (4).

Most of the rhamnose-positive *Y. pestis* isolates contain all three *Y. pestis*-specific plasmids. However, some of them lack the small pPla replicon and/or carry an enlarged pFra (8). *Y. pestis* subsp. *caucasica* (also designated Pestoides F) is frequently isolated in high mountainous Caucasus and in mountainous Dagestan. It belongs phenotypically to *Y. pestis* bv. Antiqua, and *Microtus arvalis* is its main reservoir (28). Plague epizootics of various intensities were documented in this focus. Rhamnose-positive *Y. pestis* subsp. *caucasica* strains lack pPla but contain an enlarged pFra. They have low virulence for guinea pigs. However, an aerosolized Pestoides F strain lacking the plasminogen activator was shown to be highly virulent (29). Strict geographical isolation in a high mountainous region might have led to the preservation of an ancient plague microbe. *Y. pestis* G8786, which was isolated from the high mountainous Caucasian focus, was identified as an atypical *Y. pestis* bv. Antiqua strain by genome-wide microarray analysis (11). This analysis reflected the remote origin of this organism and the highest level of divergence from other *Y. pestis* strains. Based on this knowledge, we decided to determine the whole nucleotide sequence of the enlarged pFra plasmid of rhamnose-positive *Y. pestis* strain G8786 in order to elucidate its evolutionary origin and its divergence from the pFra replicons of other *Y. pestis* isolates. The data obtained confirmed the

* Corresponding author. Mailing address: Max von Pettenkofer-Institut für Hygiene und Medizinische Mikrobiologie, Pettenkofer Str. 9a, 80336 Munich, Germany. Phone: 4989 51605275. Fax: 4989 51605223. E-mail: rakin@m3401.mpk.med.uni-muenchen.de.

chimeric origin of this plasmid (designated pG8786) and the evolutionary preservation of this potentially transmissive, ancient replicon due to strict geographical isolation.

MATERIALS AND METHODS

Bacterial strains, media, and plasmid isolation. Rhamnose-positive *Y. pestis* strain G8786 isolated from *M. arvalis* in the high mountainous Caucasus, Georgia, was a kind gift of D. Tsereteli, Georgia. This strain was cured of the pYV virulence plasmid on Luria-Bertani (LB) agar supplemented with 5 mM magnesium EGTA (Sigma, Taufkirchen, Germany) at 37°C. Loss of the pYV replicon was confirmed by plasmid screening and by PCR performed with primers YopP8 (5′-GAGACCAGTTCTTTAATCAG-3′) and YopP9 (5′-GCCAGTGCCAAAC TAAAAAT-3′) (35 cycles consisting of 30 s at 94°C, 30 s at 50°C, and 30 s at 72°C). A spontaneous Nal^r mutant of *Escherichia coli* strain JM109 (Stratagene, La Jolla, Calif.) was obtained from the strain collection of the Max von Pettenkofer Institute. Bacteria were grown at 27°C (*Y. pestis*) or 37°C (*E. coli*) in LB medium. For plasmid maintenance and mutagenesis, nalidixic acid (20 μg/ml), chloramphenicol (30 μg/ml), tetracycline (12 μg/ml), and ampicillin (100 μg/ml) were added to the culture media as required. L-Arabinose (Sigma) at a concentration of 1 mM was used for induction of the Red system genes (19) on helper plasmid pKD46. Plasmids pKD3 and pKD46 were obtained from the *E. coli* Genetic Stock Center (Yale University, New Haven, Conn.), and RP4 was obtained from the collection of the Max von Pettenkofer Institute. Plasmid DNA was isolated with a Nucleobond AX kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany).

Sequencing of pG8786. Sequencing of the pG8786 shotgun library in pUC19 was performed together with GATC Biotech AG (Konstanz, Germany). Briefly, the shotgun library was made by shearing purified pG8786 with a nebulizer. The ends of the resultant fragments were repaired with a mixture of T4 DNA polymerase and the Klenow fragment (Invitrogen, Carlsbad, Calif.). Fragments ranging from 1.2 to 2 kb long were ligated into the SmaI site of pUC19. Automated DNA sequencing was carried out by GATC Biotech AG. Sequences were assembled into contigs by using the Seqman II program (DNASTAR). The primer walking procedure was performed to close the gaps and resolve the ambiguities.

DNA sequence analysis and annotation. Open reading frames (ORFs) comprising at least 50 amino acids were identified with the Biomax Bioinformatics server (Biomax Informatics AG, Martinsried, Germany). Analysis of sequences was carried out with the BLAST program from the National Center for Biotechnology Information, the TIGR-CMR program, and Vector NTI 7.0 (InforMax).

Construction of pG8786 derivative carrying chloramphenicol resistance gene by ET mutagenesis. Construction of a pG8786 derivative carrying the chloramphenicol resistance gene was carried out as previously described (5). Briefly, electrocompetent cells were prepared from *Y. pestis* G8786 carrying the pKD46 plasmid grown in 5-ml LB medium cultures with ampicillin and L-arabinose at 27°C to an optical density at 600 nm of 0.6. G8786 competent cells were transformed with 500 to 1,000 ng of PCR product generated with PCR primers cafD.for (5′-CTGACAAATTTATGTGAAGATCAATGTTAGGAACTAATG CAGAAAGCCACGGTGTAGGCTGGAGCTGCTTC-3′) and cafD.rev (5′-A ACCCCGGGGTGAGGGCAAAGGCTGCTTTGTTGAAGTTGCATGGAT GATGGCATATGAATATCCTCCTTAG-3′) and the pKD3 plasmid as the template. Transformed cells were added to 1 ml of LB medium, incubated for 1 h at 27°C, and then spread onto LB agar to select Cm^r transformants. PCR verification was accomplished by using nearby locus-specific primers cafD1.for (5-GGGGATGACGTCGTCTTGGCTAC-3) and cafD1.rev (5-TCCACTCACT GAGTGAAGCCCTTTTAA-3) to prove correct insertion of the Cm^r cassette. Amplification of DNA by PCR was performed by using 35 cycles of 30 s at 94°C, 30 s at 60°C, and 60 s at 72°C.

Mating experiments. To determine the self-transmissivity of pG8786, mating experiments were performed on 0.45-μm-pore-size nitrocellulose filters with late-exponential-phase cultures of the donor (*Y. pestis* G8786/pG8786-Cm^r) and recipient (*E. coli* JM109 Nal^r) strains. Mating was carried out by mixing the donor and recipient strains at a ratio of 1:10 on each filter. After incubation at 27°C for 6 h on LB agar, the bacteria were plated onto selective plates. We attempted to mobilize the pG8786-Cm^r plasmid by using the conjugative RP4 plasmid that was transferred into *Y. pestis* G8786 cells as described previously (14). Subsequently, the donor *Y. pestis* G8786 cells carrying both the pG8786-Cm^r and RP4 plasmids were mated with the recipient *E. coli* JM109 Nal^r cells for 6 h at 27°C as described above.

Nucleotide sequence accession number. The annotated pG8786 nucleotide sequence has been deposited in the GenBank database under accession no. AJ698720.

RESULTS

General description. *Y. pestis* strain G8786 was cured of pYV8786 by plating on LB-EGTA agar at 37°C and selecting for loss of the Cad phenotype. Loss of the pYV8786 plasmid was proven by plasmid screening and by PCR for the pYV-encoded marker *yopP*. A shotgun library was prepared from pG8786 isolated from a *Y. pestis* G8786 monoplasmid derivative. The entire sequence of pG8786 was determined to be 137,036 bp long. Screening and annotation of the sequence with the Pedant-Pro sequence analysis suite (Biomax Informatics AG) revealed 148 putative coding regions along the entire length of the plasmid. In general, pG8786 is a pFra plasmid which has acquired *tra* genes (necessary for conjugational transfer) from an unknown source and has significant similarity to known and well-characterized conjugative plasmids, including F and R plasmids (Table 1). When a protein encoded by a pG8786 ORF had similarity to known proteins in the database, we assigned a likely function to the putative protein. A total of 62 of these ORFs are transcribed in a clockwise direction, while the remaining 86 ORFs are transcribed in a counterclockwise direction. All putative ORFs have significant homology to the genes encoding previously described hypothetical or characterized proteins in the GenBank database; 79% of them exactly match ORFs of plasmid pMT1 of *Y. pestis* bv. Mediaevalis strain 91001 (accession no. NC_005815). The positions and transcriptional orientations of all ORFs are shown in Fig. 1. In contrast to other sequenced pFra replicons (pMT1 from *Y. pestis* bv. Mediaevalis strains KIM5 and KIM10+ and pMT1 from *Y. pestis* bv. Orientalis strain CO92), this replicon contains the following two additional large coding regions: (i) three ORFs with high levels of similarity to the HCM2.0120c, HCM2.0121c, and HCM2.0122c genes of plasmid pHCM2 from *Salmonella enterica* serovar Typhi strain CT18 (designated region 1 in Fig. 2) and (ii) a large cluster of transfer genes (region 2 in Fig. 2). We did not find any crucial deletions in the pFra part of pG8786 other than the absence of two copies of the IS*100* element present in plasmid pMT1 of *Y. pestis* bv. Mediaevalis strain 91001 (Table 2).

Two potential plasmid replication regions and one partitioning system were discovered on pG8786. One replication region originated from the pFra plasmid (13, 16), while the second replication region has a high level of similarity to the alpha replicon pLV1402 plasmid of *Enterobacter intermedius* (20). The plasmid partitioning function was identical to the *parABS* system of pFra (13, 16).

The overall G+C content of pG8786 was 51.96%, compared to the lower overall G+C content of the chromosome of *Y. pestis* KIM or CO92 (47.64%) (6, 23) or of pMT1 (50.2%) (13). Surprisingly, region 2 covering nucleotides 81956 to 114573 had a G+C content of 57.58% (Fig. 2), which is much higher than the overall G+C content of the backbone of the plasmid, pointing to its horizontal acquisition.

pG8786 contains two copies of the IS*200*-like element (also known as IS*1541* in *Y. pestis*) in the opposite orientation. The first IS*200* insertion is located between positions 34588 and 35076, and the second IS*200* insertion is close to the first one (positions 36944 to 37453). Also, two copies of the IS*285*-like element were found in the opposite orientation (at positions 80712 to 81920 and 102083 to 100876). The second IS*285* had

TABLE 1. Characteristics and closest relative of the predicted product of each CDS or gene from regions 1 and 2

CDS or gene	Position (bp)	Homologue found by BLAST (ORF or gene, product, organism/element)	No. of similar amino acids/total no. of amino acids (%)	Accession no.
Region 1 (bp 37546 to 42172)				
CDS38	38438–37648	HCM2.0120c, hypothetical, protein, *S. enterica* subsp. *enterica* serovar Typhi, plasmid pHCM2	261/262 (99)	NP_569592
CDS39	39737–38595	HCM2.0121c, putative ribonucleoside diphosphate reductase beta subunit, *S. enterica* subsp. *enterica* serovar Typhi, plasmid pHCM2	379/380 (99)	NP_569593
CDS40	42160–39845	HCM2.0122c, putative ribonucleoside diphosphate reductase alpha subunit, *S. enterica* subsp. *enterica* serovar Typhi, plasmid pHCM2	769/771 (97)	NP_569594
Region 2 (bp 81956 to 114573)				
traA	82124–82480	*traA*, prepropilin, *E. coli* K-12 strain CR63, plasmid F	63/69 (91)	NP_061453
traL	82631–82936	*traL*, TraL, function in F pilus assembly, *E. coli* K-12 strain CR63, plasmid F	87/99 (87)	NP_061454
traE	82951–83517	*traE*, TraE, function in F pilus assembly, *E. coli* K-12 strain CR63, plasmid F	159/188 (84)	NP_061455
traK	83504–84247	*traK*, TraK, function in F pilus assembly, *E. coli* K-12 strain CR63, plasmid F	173/224 (77)	NP_061456
traB	84225–85628	*traB*, TraB, function in pilus assembly, *E. coli* K-12 plasmid R100-1	362/467 (77)	AAB07770
traV	85650–86189	*traV*, TraV, function in pilus biogenesis, *Shigella flexneri* 2b strain 222, plasmid R100	114/163 (69)	NP_052954
traV	86267–86533	*traR*, TraR, conjugative transfer protein, *S. enterica* serovar Typhi, plasmid pED208	55/88 (62)	AAM90708
orfX1 (CDS81)	86537–86722	*orfX1*, hypothetical protein, transfer region, *S. enterica* serovar Typhi, plasmid pED208	54/60 (90)	AAM90709
orfX2	86604–87110	*orfX2*, hypothetical protein, transfer region, *S. enterica* serovar Typhi, plasmid pED208	73/95 (76)	AAM90710
traP	87100–87666	*traP*, conjugative transfer protein, *Salmonella enterica* serovar Typhimurium LT2, plasmid pSLT	62/137 (45)	NP_490568
traC	87659–90290	Possible pseudogene, *traC, E. coli* K-12 strain CR63, plasmid F	301/339 (88)	NP_061463
trbI	90287–90634	*trbI*, pilus assembly protein, *S. enterica* serovar Typhimurium LT2, plasmid pSLT	78/91 (85)	NP_490574
traW	90631–91281	*traW*, pilus assembly protein, *S. enterica* serovar Typhimurium LT2, plasmid pSLT	143/187 (76)	NP_490575
traU	91278–92270	*traU*, pilus assembly protein, *S. enterica* serovar Typhimurium LT2, plasmid pSLT	276/320 (86)	NP_490576
trbC	92462–92917	*trbC*, hypothetical protein, *S. flexneri* 2b strain 222, plasmid R100	111/151 (73)	NP_052965
traN	92914–94764	*traN*, mating pair stabilization protein, *S. enterica* serovar Typhimurium LT2, plasmid pSLT	455/607 (74)	NP_490579
traF	95017–95766	*traF*, TraF, *E. coli*, plasmid R100-1	193/234 (82)	AAB61943
traO	95693–96031	*traQ*, pilin chaperone, *S. enterica* serovar Typhimurium LT2, plasmid pSLT	47/63 (74)	NP_490582
trbB	95970–96566	*trbB*, hypothetical protein, *E. coli* K-12 strain CR63, plasmid F	139/179 (77)	NP_061474
traH	96563–97936	*traH*, pilus assembly protein, *S. enterica* serovar Typhimurium LT2, plasmid pSLT	394/446 (88)	NP_490584
traG	97936–100755	*traG*, responsible for pilus biogenesis and stabilization of mating pairs, *S. flexneri* 2b strain 222, plasmid R100	761/939 (81)	NP_052976
Y1062	102083–100876	Y1062, possible pseudogene, putative transposase IS*285*, *Y. pestis* KIM10+, plasmid pMT-1	381/381 (100)	AAC82722
traD	102087–104276	*traD*, TraD, *E. coli* F sex factor	547/705 (77)	AAC44181
traI	104276–109516	*traI*, TraI, function in *oriT* nicking and unwinding, *E. coli* K-12 strain CR63, plasmid F	1,324/1,759 (75)	NP_061483
traX	109380–110255	*traX*, F pilin acetylation protein, *S. flexneri*, plasmid pWR501	160/284 (56)	NP_085416
finO	110322–111041	*finO*, fertility inhibition protein, *E. coli*, plasmid F	75/180 (41)	P22707
bcfH	111028–111825	*bcfH*, putative thiol disulfide isomerase, *S. enterica* serovar Typhimurium LT2	138/244 (56)	NP_459033
nuc	111973–112527	*nuc*, endonuclease, *Y. enterocolitica* strain 15673 (biogroup 1A serogroup 0:5) pYV plasmid	135/174 (77)	CAA73744
copB	112418–112648	*copB*, CopB, function in replication control, *E. intermedius*, plasmid pLV1402	36/46 (78)	CAA08928
copA	112835–112750	*copA*, small RNA, function in antisense control, *E. coli*, plasmid RI	71/81 (87)	V00326
tapA	112867–112944	*tapA*, TapA, function in antisense control, *E. intermedius*, plasmid pLV1402	24/25 (96)	CAA08929
repA	112925–113800	*repA*, replication protein, *E. intermedius*, plasmid pLV1402	234/287 (81)	CAA08930
SMR0139	114351–114847	*SMR0139*, possible pseudogene, ATP/GTP-binding protein, *Serratia marcescens*, plasmid R478	143/164 (87)	NP_941209

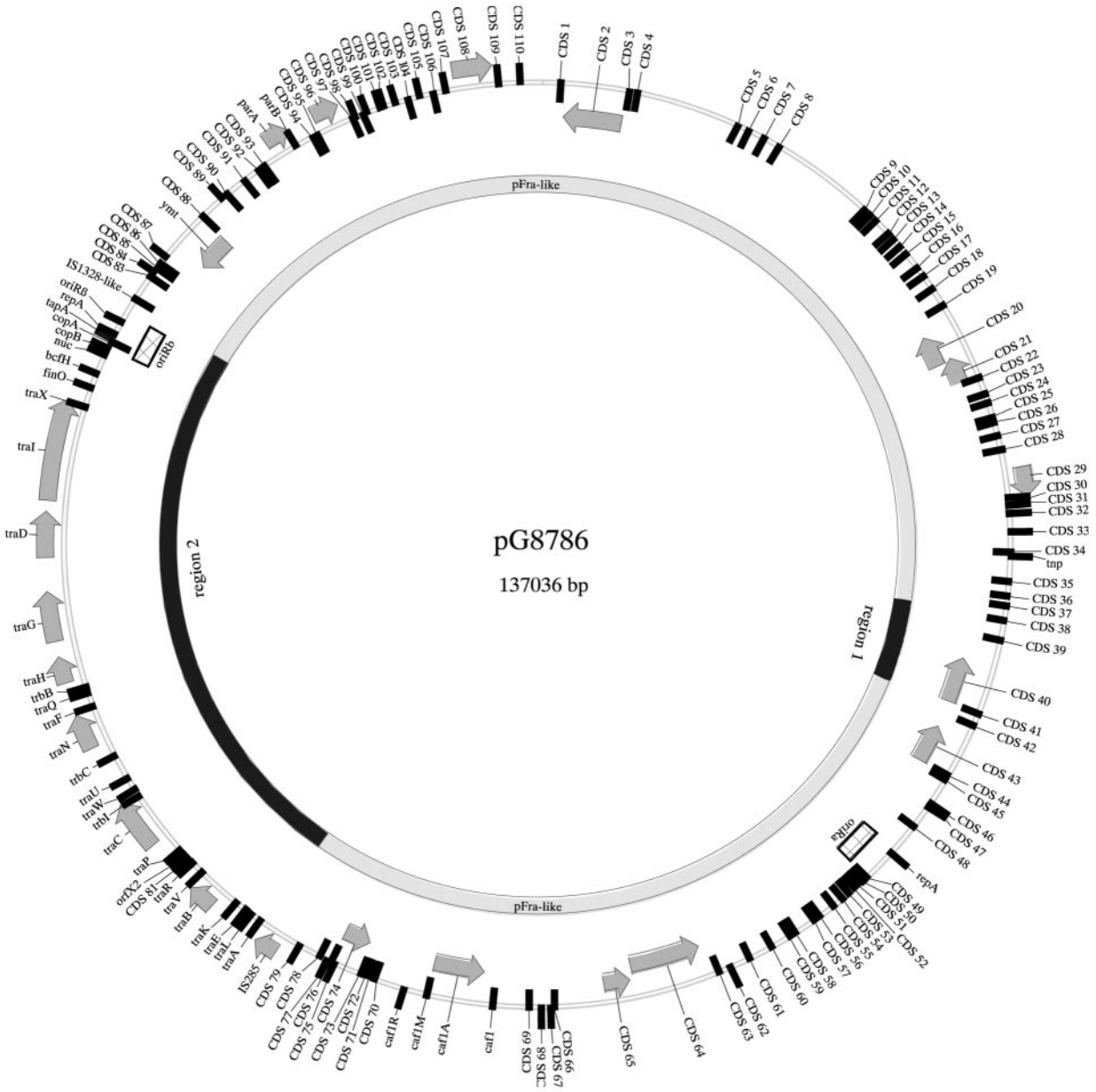

FIG. 1. Map of the pG8786 plasmid. The inner circle shows region 1, region 2, and the pFra-like backbone. The outer circle shows ORFs and their orientation, which are designated on the basis of their positions; the arrows and boxes outside the ring indicate clockwise transcription, and the arrows and boxes inside the ring indicate counterclockwise transcription. The map was derived from the annotated DNA sequence by using the Vector NTI (InforMax) computer program and was edited in CorelDRAW.

a frameshift after codon 169 and thus appeared to be a nonfunctional remnant. One copy of each element, an IS*1328*-like element and an IS*100* element, was found on pG8686 (Table 2 and Fig. 2). These copies had the same orientation as the first copy of IS*285*.

ORFs of region 1. The 4,626-bp region 1 is not present in the pFra plasmids of *Y. pestis* KIM and CO92 representing *Y. pestis* bv. Mediaevalis and *Y. pestis* bv. Orientalis, respectively (Fig. 3). Surprisingly this region is present in the pFra plasmid of avirulent *Y. pestis* bv. Mediaevalis strain 91001 isolated from *Microtus brandti* in Inner Mongolia, People's Republic of China (accession no. AE017045). It is also 96% identical to plasmid pHCM2 of *S. enterica* serovar Typhi strain CT18 (22). Our analysis revealed three putative ORFs (CDS38, CDS39, and CDS40) spanning bp 37641 to 42160 (Table 1). The CDS38 protein is very similar to the HCM2.0120c hypothetical protein. The CDS39 and CDS40 proteins are putative beta and alpha ribonucleoside diphosphate reductase subunits which might be necessary for deoxyribonucleotide metabolism.

ORFs of the transfer region. ORFs spanning bp 81956 to 111038 (region 2) in pG8786 were found to be similar to the *tra* region genes of the F-like plasmids which belong to the type IV family of secretion systems (Table 1). Analysis of the pG8786 *tra* region revealed 25 putative ORFs (*traA-bcfH*), whereas the

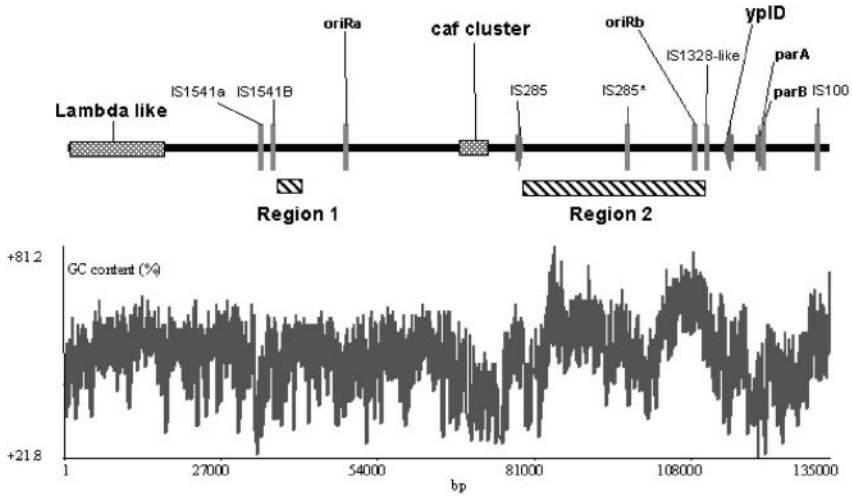

FIG. 2. G+C content and graphic map of pG8786. The plot showing the G+C content was derived by using the Vector NTI program (InforMax). The diagram at the top shows selected ORFs and some other annotated features at the correct scale. The scale below the G+C plot indicates the size of the plasmid. IS285*, IS*285* insertion sequence which appeared to be a nonfunctional remnant.

F *tra* region has 37 ORFs (9). However, the cryptic conjugative plasmid from *Yersinia enterocolitica* 29930 contains even fewer genes (i.e., 16 ORFs) (27). The *tra* region of pG8786 is organized like and highly homologous to the *tra* regions of F-like plasmids belonging to different incompatibility groups of the IncF family, including IncFI (F), IncFII (R-100, R100-1), and IncFV (pED208) (Table 1 and Fig. 4a). Accordingly, the putative pG8786 transfer genes were designated on the basis of their homologs in the IncF plasmids.

The gene products can be organized into the following four groups based on functions inferred from their closely related homologs: pilus biogenesis (TraA-V, TraW, TraU, TrbC, TrbI, TraF, TraH, TraQ, TraX, and the N-terminal region of TraG); regulation (FinO); DNA nicking and initiation of transfer (TraI and TraD); and mating aggregate stabilization (TraN and TraG). Other nonclassified components encoded in the *tra* region are TraP, a protein that stabilizes the extended pilus; TrbB, a putative thioredoxin homolog; and hypothetical proteins OrfX1 and OrfX2 (9, 15, 17).

Surprisingly, genes with similarity to *traM* (whose function is signaling that DNA transfer should begin), *traJ* (a positive regulator of transcription of the *tra* operon), *traT*, and *traS* (surface exclusion) were not detected in the *tra* region of pG8786. We also identified only the 3′-terminal remnant of *traY* located next to the first copy of IS*285*. Since the gene organization of the *tra* region of pG8786 most closely resembles the organization of the transfer region of plasmid pED208 of *S. enterica* serovar Typhi (Fig. 4a and 5) (17), we speculate that the *traM*, *traJ*, *traY*, *traT*, and *traS* genes might be deleted or truncated (in the case of *traY*).

The *oriT* region is arbitrarily defined as the region at the beginning of the *traM* gene. This region contains the site where nicking occurs and transfer of the single-stranded DNA, in a 5′-to-3′ manner, into the recipient cell is initiated (9). As the *traM-Y* gene locus was absent from pG8786, we tried to define a possible origin of transfer (*oriT*). However, an expanded search of the pG8786 sequence did not reveal any region of the plasmid that might function as *oriT*. Nevertheless, plasmids without a defined *oriT* have been described (10).

An interesting feature of the pG8786 *tra* region is the presence of the gene corresponding to *finO* (Table 1). FinO is a part of the FinOP system that is a key determinant defining the frequency of IncF plasmid-mediated DNA transfer. TraJ, a positive regulator of the *tra* genes, is controlled at the post-transcriptional level by two negative elements, *finP* and *finO*. FinP is a plasmid-specific antisense RNA, whereas *finO* encodes a corepressor, which is almost identical to (level of identity, more than 95%) and cross-reactive with various F-like plasmids (9). We hypothesize that the intact FinO of pG8786 can also repress transcription of *traJ* of other F-like conjugative plasmids which might have been acquired by G8786 cells. Thus, the frequency of transfer of such an acquired plasmid might be dramatically reduced.

To check the self-transmissivity of pG8786, we inserted a

TABLE 2. Distribution of insertion elements in five sequenced pFra plasmids of *Y. pestis*

Y. pestis strain	No. of IS*100* copies	No. of IS*285* copies	No. of IS*1541* (IS*200*) copies	No. of IS*1328*-like copies	Accession no.
91001	3	1	1	1	NC_005815
KIM10+	2	1	1	1	AF074611
KIM5	2	1	1	1 (IS*1618*)	AF053947
CO92	2	1	1	0	NC_003134
G8786	1	2	2	1	AJ698720

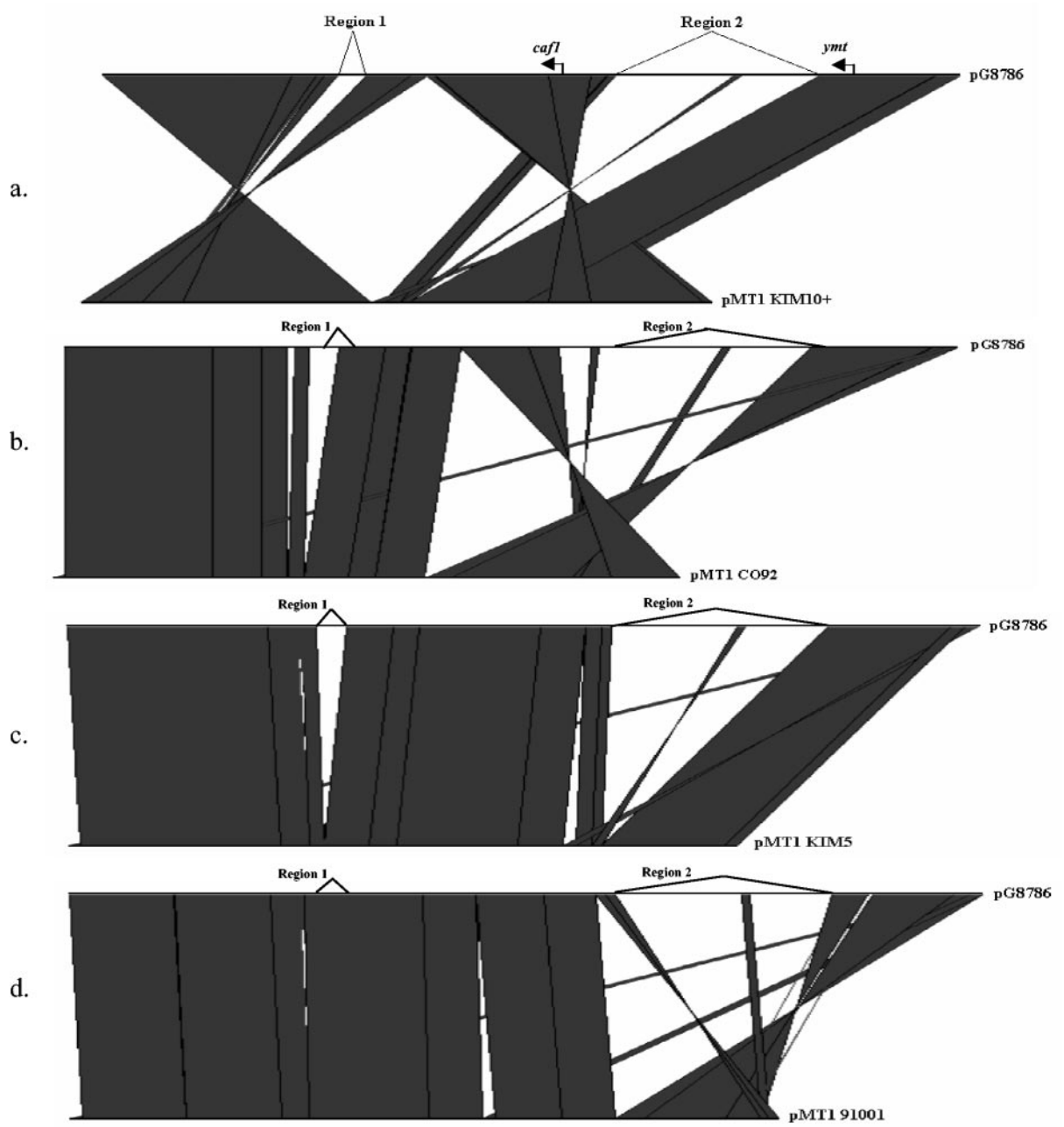

FIG. 3. Graphic comparison of different pFra plasmids with pG8786 derived by using the Artemis Comparison Tool program (The Wellcome Trust Sanger Institute, Cambridge, United Kingdom). (a) pG8786 and pMT-1 of *Y. pestis* KIM10+; (b) pG8786 and pMT1 of *Y. pestis* CO92; (c) pG8786 and pMT1 of *Y. pestis* KIM5; (d) pG8786 and pMT1 of *Y. pestis* 91001. Areas of pG8786 that are not present in the other pFra plasmids are labeled region 1 and region 2. Vertical lines indicate similar parts of the plasmids.

chloramphenicol gene cassette between the *caf1* gene and CDS69 (primer positions 69955 to 70412) to tag this plasmid. *Y. pestis* G8786(pG8786-Cm^r) was mated with the recipient *E. coli* JM109 (Nal^r). However, we did not detect transfer of the Cm^r marker and thus of the labeled plasmid. Also, our attempts to mobilize pG8786-Cm^r with the broad-host-range RP4 IncP-alpha plasmid (21) were unsuccessful. Nevertheless, we cannot exclude the possibility that pG8786 might be transmissible if it is supplemented with the missing *tra* genes in *trans* or if it is mated with a more suitable recipient strain.

Replication and plasmid maintenance. As mentioned above, DNA sequence analysis revealed two potential origins of plasmid replication which were designated *oriRa* (bp 49713 to 51622) and *oriRb* (bp 112331 to 114155) (Fig. 1 and 2). *oriRa* is identical to the replication origin of pFra, which is also similar to RepFIB, RepHI1B, and the P1 and P7 replicons (15).

The second origin of replication, *oriRb*, which was localized in region 2, showed very high similarity (89%) to the alpha replicon (RepFIIA) of pLV1402 of *E. intermedius* (20). It is

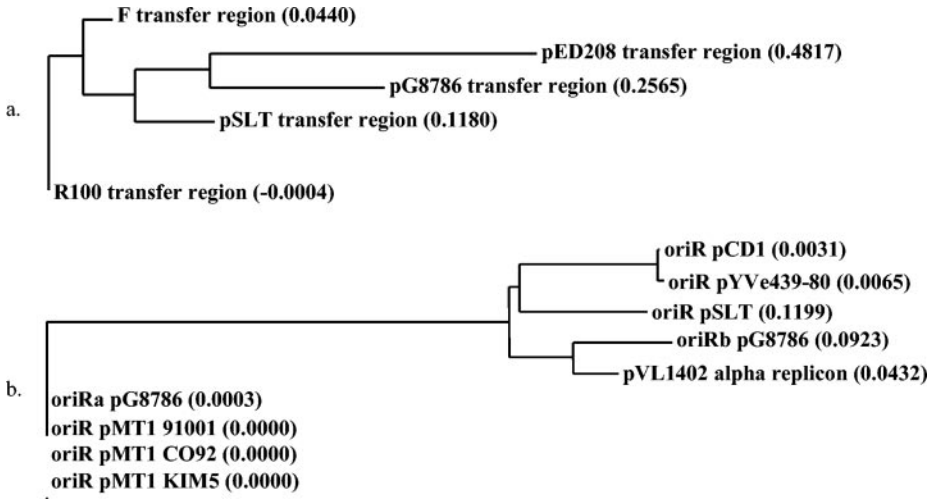

FIG. 4. (a) Phylogenetic tree for the transfer regions from various sources. (b) Phylogenetic tree for the replication origins from different microorganisms. The trees were constructed by using the neighbor-joining method of Saitou and Nei and the Vector NTI program (InforMax). Using AlignX, we determined the calculated distance values (which are indicated in parentheses following the molecule designations on the trees).

closely related to the IncFII virulence-associated replicons of pCD1 of *Y. pestis* (accession no. AF074612) and pYVe439-80 of *Y. enterocolitica* (accession no. M55182) (Fig. 4b). On the basis of the similarity to these replicons, the following genetic features were identified: *copB* (bp 112331 to 112648), *copA* (bp 112835 to 112750), *tapA* (bp 112867 to 112944), *repA* (bp 112925 to 113800), and *oriR*β (bp 113972 to 114155) (Table 1 and Fig. 1). In the IncFII replicons an antisense RNA molecule (CopA) inhibits synthesis of the replication protein (RepA) by binding to the leader region of the *repA* mRNA (CopT). RepA synthesis depends on translation of a short leader peptide (TapA) that is not expressed when CopA binds to CopT, thereby preventing translation of RepA and consequently preventing replication of the plasmid (3). The sequence 5′-TTG CCCACA-3′, which may function as a binding site for the DnaA protein, could be defined 174 bp downstream of *repA*.

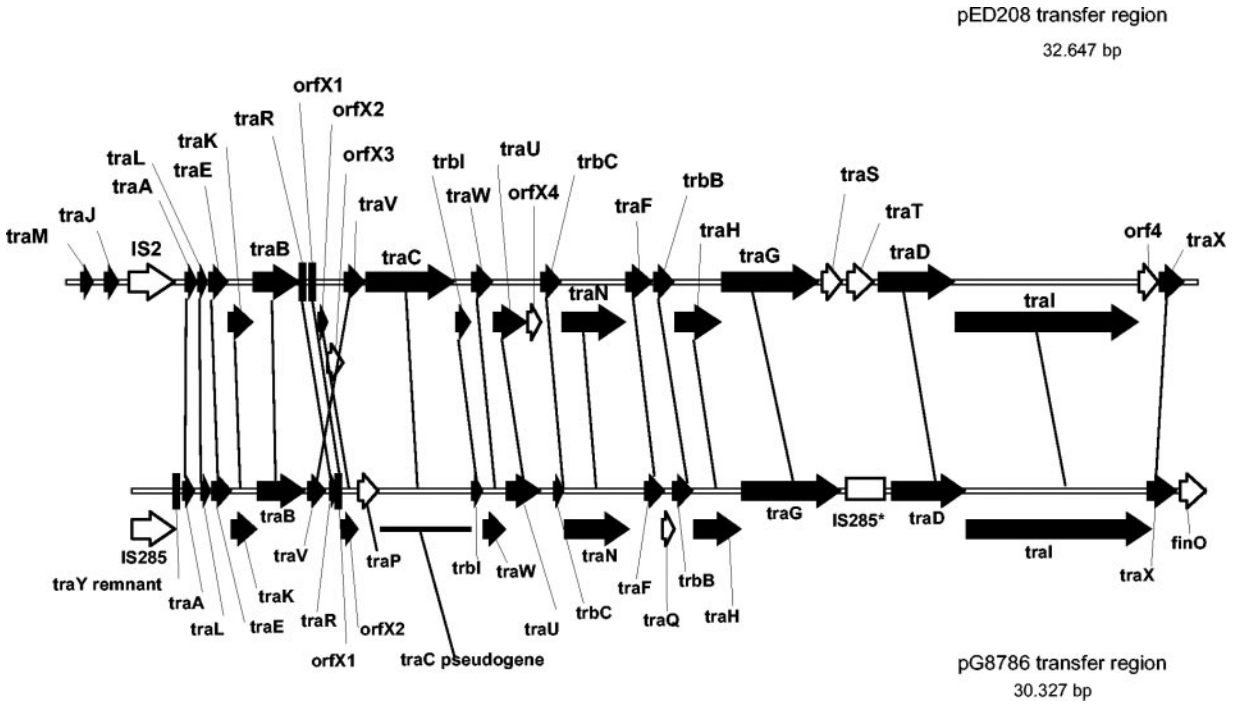

FIG. 5. Alignment of the *tra* genes of pG8786 (lower diagram) with the *tra* genes of the pED208 plasmid from *S. enterica* serovar Typhi (upper diagram). The solid arrows represent ORFs which are equally represented in the two plasmids. The open arrows represent ORFs which are absent in either of the two transfer regions. IS285*, IS*285* insertion sequence which appeared to be a nonfunctional remnant. The vertical lines between arrows indicate similar ORFs.

This sequence matches the DnaA box at seven of nine positions (24).

DISCUSSION

The complete sequence of pG8786, the 137-kb virulence plasmid of rhamnose-positive *Y. pestis* strain G8786 obtained from a vole (*M. arvalis*) in a locus in the high mountainous Caucasus region in Georgia where plague is endemic, was determined. This sequence revealed the recombinant nature of pG8786 (namely, insertion of a 32,617-bp unique *tra* gene cluster and *oriRb* [region 2]). Also, a larger part of plasmid pHCM2 of *S. enterica* serovar Typhi that forms the backbone of the *Y. pestis*-specific plasmid pFra (25) is preserved in pG8786 (region 1). However, in contrast to other pFra plasmids (pMT1 from *Y. pestis* bv. Orientalis strain CO92 and pMT1 from *Y. pestis* bv. Mediaevalis strain KIM), the complete region 1 was also found in the pFra plasmid of avirulent *Y. pestis* bv. Mediaevails strain 91001 isolated from another vole (*M. brandti*) in Inner Mongolia, People's Republic of China. The fact that a larger remnant of pHCM2, which is the same size, is present in both geographically isolated atypical *Y. pestis* strains but is not present the epidemic isolates implies that plasmid pFra in *Y. pestis* G8786 and 91001 might represent an ancient form of pFra. This hypothesis is supported by the presence of an additional transfer region in pG8786. Alternatively, pG8786 acquired the *tra* operon more recently by horizontal gene transfer. The variations in the G+C content also point to the chimeric nature of the plasmid. With these facts in mind, one can suppose that pG8786 may have originated by acquisition of DNA fragments from various microorganisms with higher G+C contents.

A comparison of the four sequenced pFra plasmids with pG8786 revealed extended regions of DNA rearrangements in the backbone of the pFra replicon (i.e., large inversions resulting from insertion element-mediated recombination). Thus, the presence of these flanking insertion elements might reveal DNA rearrangements in the pG8786 plasmid. We hypothesize that pFra initially cointegrated with a conjugative IncFII group plasmid and that this was followed by subsequent deletion by IS*285*-mediated recombination. Certain *tra*-associated genes (*traM*, *traJ*, *traY*, *traT*, *traS*, and *par*) of pG8786 or the complete *tra* region of another pFra plasmid (pMT1 91001) have such a deletion. Alternatively, the presence of the *par* partition genes, associated with the second RepFII replication origin, might decrease the ability of pG8786 to coexist with the virulence-associated pYV replicon that probably belongs to the same incompatibility group.

Conjugative transfer of the Cm^r-labeled pG8786 plasmid could not be demonstrated by using the original host, *Y. pestis* G8786, and *E. coli* as the recipient. Nevertheless, genetic exchange might take place under certain conditions, possibly in a different ecological environment (e.g., the flea midgut) (12). Acquisition of a transmissive form of the pFra plasmid (encoding the fraction 1 antigen and phospholipase D necessary for colonization of the flea gut) might be the first major step in *Y. pestis* evolution from a common ancestor of *Y. pseudotuberculosis* and *Y. pestis*. Such an acquisition, combined with the preexisting sequences in the genomes of pathogenic yersiniae (*Y. pseudotuberculosis* [http://bbrp.llnl.gov/bbrp/html/microbe.html] and *Y. enterocolitica* [http://www.sanger.ac.uk/Projects/Y_enterocolitica/]) for the pigmentation locus and insect toxin genes, provided the evolving organisms with a greater ability to survive in the flea vector, resulting in efficient blood-borne transmission. The existence of a potentially transmissive virulence-associated plasmid in *Y. pestis* points to the fact that occasionally a new pathogen may appear with the ability to survive and multiply efficiently in insect vectors like fleas.

REFERENCES

1. **Achtman, M., K. Zurth, G. Morelli, G. Torrea, A. Guiyoule, and E. Carniel.** 1999. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. Proc. Natl. Acad. Sci. USA **96:**14043–14048.
2. **Anisimov, A. P., L. E. Lindler, and G. B. Pier.** 2004. Intraspecific diversity of *Yersinia pestis*. Clin. Microbiol. Rev. **17:**434–464.
3. **Blomberg, P., K. Nordstrom, and E. G. Wagner.** 1992. Replication control of plasmid R1: RepA synthesis is regulated by CopA RNA through inhibition of leader peptide translation. EMBO J. **11:**2675–2683.
4. **Cornelis, G. R., and H. Wolf-Watz.** 1997. The *Yersinia* Yop virulon: a bacterial system for subverting eukaryotic cells. Mol. Microbiol. **23:**861–867.
5. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA **97:**6640–6645.
6. **Deng, W., V. Burland, G. Plunkett III, A. Boutin, G. F. Mayhew, P. Liss, N. T. Perna, D. J. Rose, B. Mau, S. Zhou, D. C. Schwartz, J. D. Fetherston, L. E. Lindler, R. R. Brubaker, G. V. Plano, S. C. Straley, K. A. McDonough, M. L. Nilles, J. S. Matson, F. R. Blattner, and R. D. Perry.** 2002. Genome sequence of *Yersinia pestis* KIM. J. Bacteriol. **184:**4601–4611.
7. **Devignat, R.** 1951. Variétés de l'espèce *Pasteurella pestis*: nouvelle hyphothèse. Bull. W.H.O.**4:**-263.
8. **Filippov, A. A., N. S. Solodovnikov, L. M. Kookleva, and O. A. Protsenko.** 1990. Plasmid content in *Yersinia pestis* strains of different origin. FEMS Microbiol. Lett. **67:**45–48.
9. **Frost, L. S., K. Ippen-Ihler, and R. A. Skurray.** 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol. Rev. **58:**162–210.
10. **Galli, D. M., J. Chen, K. F. Novak, and D. J. Leblanc.** 2001. Nucleotide sequence and analysis of conjugative plasmid pVT745. J. Bacteriol. **183:**1585–1594.
11. **Hinchliffe, S. J., K. E. Isherwood, R. A. Stabler, M. B. Prentice, A. Rakin, R. A. Nichols, P. C. Oyston, J. Hinds, R. W. Titball, and B. W. Wren.** 2003. Application of DNA microarrays to study the evolutionary genomics of *Yersinia pestis* and *Yersinia pseudotuberculosis*. Genome Res. **13:**2018–2029.
12. **Hinnebusch, B. J., A. E. Rudolph, P. Cherepanov, J. E. Dixon, T. G. Schwan, and A. Forsberg.** 2002. Role of *Yersinia* murine toxin in survival of *Yersinia pestis* in the midgut of the flea vector. Science **296:**733–735.
13. **Hu, P., J. Elliott, P. McCready, E. Skowronski, J. Garnes, A. Kobayashi, R. R. Brubaker, and E. Garcia.** 1998. Structural organization of virulence-associated plasmids of *Yersinia pestis*. J. Bacteriol. **180:**5192–5202.
14. **Kim, S. R., N. Funayama, and T. Komano.** 1993. Nucleotide sequence and characterization of the *traABCD* region of IncI1 plasmid R64. J. Bacteriol. **175:**5035–5042.
15. **Lawley, T. D., W. A. Klimke, M. J. Gubbins, and L. S. Frost.** 2003. F factor conjugation is a true type IV secretion system. FEMS Microbiol. Lett. **224:**1–15.
16. **Lindler, L. E., G. V. Plano, V. Burland, G. F. Mayhew, and F. R. Blattner.** 1998. Complete DNA sequence and detailed analysis of the *Yersinia pestis* KIM5 plasmid encoding murine toxin and capsular antigen. Infect. Immun. **66:**5731–5742.
17. **Lu, J., J. Manchak, W. Klimke, C. Davidson, N. Firth, R. A. Skurray, and L. S. Frost.** 2002. Analysis and characterization of the IncFV plasmid pED208 transfer region. Plasmid **48:**24–37.
18. **Martinevskii, I. L.** 1969. Biology and genetic features of plague and plague-related microbes. Meditsina Press, Moscow, USSR.
19. **Murphy, K. C.** 1998. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. J. Bacteriol. **180:**2063–2071.
20. **Osborn, A. M., F. M. da Silva Tatley, L. M. Steyn, R. W. Pickup, and J. R. Saunders.** 2000. Mosaic plasmids and mosaic replicons: evolutionary lessons from the analysis of genetic diversity in IncFII-related replicons. Microbiology **146:**2267–2275.
21. **Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas.** 1994. Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. J. Mol. Biol. **239:**623–663.
22. **Parkhill, J., G. Dougan, K. D. James, N. R. Thomson, D. Pickard, J. Wain, C. Churcher, K. L. Mungall, S. D. Bentley, M. T. Holden, M. Sebaihia, S. Baker, D. Basham, K. Brooks, T. Chillingworth, P. Connerton, A. Cronin, P. Davis, R. M. Davies, L. Dowd, N. White, J. Farrar, T. Feltwell, N. Hamlin,**

A. Haque, T. T. Hien, S. Holroyd, K. Jagels, A. Krogh, T. S. Larsen, S. Leather, S. Moule, P. O'Gaora, C. Parry, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead, and B. G. Barrell. 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature **413:**848–852.

23. **Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. Holden, M. B. Prentice, M. Sebaihia, K. D. James, C. Churcher, K. L. Mungall, S. Baker, D. Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T. Chillingworth, A. Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Leather, S. Moule, P. C. Oyston, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead, and B. G. Barrell.** 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague. Nature **413:**523–527.
24. **Praszkier, J., T. Wei, K. Siemering, and J. Pittard.** 1991. Comparative analysis of the replication regions of IncB, IncK, and IncZ plasmids. J. Bacteriol. **173:**2393–2397.
25. **Prentice, M. B., K. D. James, J. Parkhill, S. G. Baker, K. Stevens, M. N. Simmonds, K. L. Mungall, C. Churcher, P. C. Oyston, R. W. Titball, B. W. Wren, J. Wain, D. Pickard, T. T. Hien, J. J. Farrar, and G. Dougan.** 2001. *Yersinia pestis* pFra shows biovar-specific differences and recent common ancestry with a *Salmonella enterica* serovar Typhi plasmid. J. Bacteriol. **183:** 2586–2594.
26. **Skurnik, M., A. Peippo, and E. Ervela.** 2000. Characterization of the O-antigen gene clusters of *Yersinia pseudotuberculosis* and the cryptic O-antigen gene cluster of *Yersinia pestis* shows that the plague bacillus is most closely related to and has evolved from *Y. pseudotuberculosis* serotype O:1b. Mol. Microbiol. **37:**316–330.
27. **Strauch, E., G. Goelz, D. Knabner, A. Konietzny, E. Lanka, and B. Appel.** 2003. A cryptic plasmid of *Yersinia enterocolitica* encodes a conjugative transfer system related to the regions of CloDF13 Mob and IncX Pil. Microbiology **149:**2829–2845.
28. **Tsereteli, D., L. Bakanidze, I. Velijanashvili, L. Beridze, M. Manrikyan, M. Shakhikyan, M. Kekelidze, E. Zangaladze, N. Tsertsvadze, and P. Imnadze.** 2002. Plague in Southern Caucasus. Antibiot. Monitor **18:**14–16.
29. **Worsham, P. L., and C. Roy.** 2003. Pestoides F, a *Yersinia pestis* strain lacking plasminogen activator, is virulent by the aerosol route. Adv. Exp. Med. Biol. **529:**129–131.

Editor: J. B. Bliska